Journal of Current Research in Engineering and Science
Bi-Annual Online Journal ( ISSN : 2581 - 611X )
[Type here]

Volume 6- Issue 2, August 2023
Paper : 34

# Survey on Aspect-Based Sentiment Analysis

Jitendra Soni*, Dr. Kirti Mathur**, Mr Arpit Agrawal *, Karan Bais *, Pravar More*, Harsh Gehlot *
*Institute of Engineering & Technology, Devi Ahilya VishwaVidyalaya, Indore
**International Institute of Professional Studies, Devi Ahilya Vishwavidyalaya, Indore

Corresponding author's e-mail: jitendrasoni2017@gmail.com

## Abstract

Sentiment analysis is the process of determining the Sentiment depicted in a statement by the author. Sentiment analysis has been a growing field since the last decade and with the ever-increasing demand of companies and individuals to understand the needs and opinions of their audience has led the field to go the extra mile by introducing many different algorithms and techniques to achieve more correct and informative conclusions about their offerings. The field has been subdivided into different branches each addressing a different level of analysis and ideas. The following survey focuses on the current work in this field, especially the new idea of aspect-based sentiment analysis, which focuses on finding the individual sentiment about entities discussed in the sentence and document. The importance of Aspect Based Sentiment analysis lies in its ability to yield much more fine-grained sentiment information than its peers which is discussed in detail. The paper goes into detail about the works done on Aspect Based Sentiment analysis in the near past and breaks down the exact motives and features of these works, along with describing the proposed solutions in detail and comparing their performance based on the standard metrics. Finally, the paper also discusses the challenges and problems faced in the task and expands on what future work can be done to tackle them.

**Keywords:** Sentiment Analysis, Aspect-Based Sentiment Analysis, Natural Language Processing

## 1 Introduction

With the onset of the digital age, information has become a hot commodity that is used in almost all industries. With the ongoing developments in the World Wide Web, the volume of user-generated data has been growing rapidly. Along with this newly generated data, several opinions about different topics have emerged each showing a different perspective on these topics. These preferences present an opportunity for the companies and service providers to better understand the needs of consumers and allow them to fine-tune their products to ful fill them. Thus, this new information processing and data mining field has gained popularity due to its widespread applications.

### 1.1 Applications

For the longest time, many organizations have faced problems when it comes to the opinions and preferences their audience has on their

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal ( ISSN : 2581 - 611X )

Volume 6- Issue 2, August 2023
Paper : 34

products and services as this would allow them to make the decisions in the right direction and make more personalized products for their audience. This challenge was due to the lack of data repositories and tedious collection methods. With the rise of the internet and multiple places where people can express their opinions about different things, the issue of lack of data is solved to some extent. Now, opinions about any subject can be procured in real-time, allowing researchers to focus on other problems.

Sentiment Analysis can be used as Decision-making support. Using data from user reviews, and reactions to potential change updates on any social media platform in the form of comments, etc. Sentiment analysis can help producers make informed decisions about how they should proceed. Sentiment scrutiny can also be used to predict upcoming trends and the direction the current trends may or may not take.

**1.2 Definitions**

Sentiment analysis has been called many different names due to its operation domain, for example, it is called opinion mining from the context of data mining and information retrieval and focuses on determining the opinion on a topic. The term sentiment analysis focuses on the idea of deducing the sentiment based on textual information. However, all of these fields focus on the same idea "the study of phenomena of opinion, sentiment, evaluation, appraisal, attitude, and emotion" [1].

A statement can be classified into two different pairs of classes which are mutually independent but can collectively define major features of the statement. These features are – Is the statement subjective or objective and whether it contains a sentiment or not? A subjective statement suggests the personal idea of an individual toward a topic. For example, "To me, the car looks sportier than a family car." And "I like a family car look over a sporty look." Both statements suggest a personal or subjective opinion on a topic but while the first statement shows no Sentiment toward the topic the second statement suggests a clear sentiment. Also, "Consensus classifies the Car as an economical choice considering the current conditions." And "The idea of an economical car has been a more agreeable option against a more feature-laden car" shows the example of two separate statements each showcasing an objective opinion on a topic but while the first was lacking a sentiment the second shows a clear one. This idea can be extended further to a more fine-grained idea of Sentiment by thinking about the idea of what part of the sentence is the sentiment more focused on i.e., the "aspect" that the Sentiment is about. So, the sentences "The idea of an economical car has been a more agreeable option against a more feature-laden car" and "I like a family car look over a sporty look." both have shown a separate sentiment about two separate parts of the sentences. In the first sentence, the aspect of the economical car received a positive sentiment and the "feature-laden" car received a comparatively negative sentiment. Similarly, the author prefers the idea of a car with a "family looks" to a car with a "sporty look". With this idea in mind, Sentiment detection can be defined as trying to discover the quadruple (Sentiment, target object, sentiment descriptor, time of expression) [1]. However, the paper discussed in the following survey focuses mainly on the first two values of these quadruples and it is mainly dependent on the application domain of these works.

**1.3 The focus of the survey**

The focus of this survey is to discuss the current works being done in the domain of Aspect-based Sentiment analysis. Sentiment analysis is classified, based on what level of the document is being analyzed as the smallest unit, into three different parts –

- Document-based Sentiment analysis – we consider the entire document as a whole and try to decide what the author's sentiment is in the document.
- Sentence-based Sentiment analysis – Here the document is broken down into separate Sentences and then analyzed one by one to discover what sentiment each sentence is portraying.

- Aspect-based Sentiment analysis – The sentence is further broken down into two parts, the aspect or the target of a sentiment. It can be an object, person, or any entity per se. And the second part is the Sentiment of the aspect. Each Sentence can contain zero to multiple aspects and build a fine-grained corpus of these aspects and define what is their

Journal of Current Research in Engineering and Science
[Type here]
JCRES
Bi-Annual Online Journal ( ISSN : 2581 - 611X )

**Volume 6- Issue 2, August 2023**
**Paper : 34**

overall sentiment in the document as an aggregate.

The reason for us to focus on this topic is due to its rising popularity amongst researchers and its proven quality in the field of Sentiment analysis. With the ability of aspect-level sentiment analysis to be more concise with its sentiment indication, it has naturally piqued the interest of many researchers.

Many surveys have been done on the topic of Aspect-based Sentiment analysis to describe what new work has been going on in the field to update interested individuals on the progress. In the following paragraph, we discuss some of them.

### 1.4 Literature survey

One of the first surveys, done by B. Liu [1], discuss in detail about the three levels of sentiment classification, document, sentence, and the entity or aspect level Sentiment analysis in detail and showcase multiple different application of the same in the form of practical works like "Opinion Search and Retrieval", "Opinion Spam Detection", "Quality of Review", to showcase the work in Sentiment analysis. The chapter that deals with Aspect Based Sentiment analysis discusses the problems that we face in the case of Sentiment analysis and talks about two major problems – extracting or identifying the aspect and classifying the sentiment polarity based on context words. The algorithms he states show a bad accuracy due to their inability to deal with complex statements and handle factual information that conveys a sentiment in an implicit way [1]. Another study done by K. Schouten and F. Frasincar [3] goes into classifies the work done during that time based on their work methodology. The paper divides the task of Aspect-Based Sentiment Classification into three independent tasks – Aspect Detection, Sentiment Analysis, and Joint Aspect Detection and Sentiment Analysis and divide these works based on which task they undertake. Furthermore, the paper goes into subcategories of these tasks dividing them on the technique being used in these tasks, and further subclassifies the papers in discussion in those categories. Also, they discuss the efforts being taken in the field to standardize the method of evaluation for Sentiment Classification. Finally, they discuss the Related issues that Sentiment analysis entails and discuss how to aggregate and present the findings of such projects. Another major survey that was done in the field of Sentiment analysis was done by A. Nazir et. al. [4]. The survey focuses on providing a summary of papers considered from as far back as 2016 and discusses different parts of Sentiment analysis and the way the field has grown in the last few years. The paper discusses the works that focused on Aspect Extraction and multiple different methods of doing the same for example Implicit Aspect Extraction, Extraction with neutral Sentiment, Cross Domain, and Cross-Lingual Aspect extraction. Then it discusses why it is essential to map the relationship between certain aspects and relations and how to do so (via Co-occurrence relationship, Sematic Relationship, etc.), followed by a discussion of different ways to perform Aspect Sentiment analysis via methods of Target-level, Entity-level, multi-Task and state the relevant works that were done on the same in the last few years. The methods usable for further improving the process are also discussed (like Data object Interactions). Finally, they discuss the factors that cause Sentiment evolution that causes changes in opinions about the aspects over time finishing off with recommendations for future works.

# 2 Evaluation Methodology

### 2.1 Evaluation Measures

Evaluation metrics are used to judge how the model is performing and give a different perspective of a model's performance. Thus, choosing the right evaluation methods is essential. Each provides a different insight and should be chosen as per the model's application needs. Most prominently Accuracy, F1 Score, and recall are used. In addition, various measures such as Ranking Loss, Mean Absolute Error, and Mean Squared Error (MSE) are utilized.
**Accuracy**: This represents the measure of how many correct predictions were made based on the ratio of correct predictions to the total predictions. Now, accuracy is a flawed measure when we consider a case where the dataset is

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal ( ISSN : 2581 - 611X )


**Volume 6- Issue 2, August 2023**
**Paper : 34**

imbalanced as it may lead to a model which is not capable of generalization.

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN}$$

**Precision**: it is defined as the tendency of the model to be able to make correct predictions. It measures the quality of correct predictions the model is making

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$= \frac{True\ Positive}{Total\ Predicted\ Positive}$$

In simple language let our model predict that out of 10 sentence 7 has positive sentiment and among that predicted 7, only 3 were positive, so in this case precision is 3/7 = 0.428.

**Recall**: Recall calculates how well the model can recall the template of correct predictions.

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$= \frac{True\ Positive}{Total\ Actual\ Positive}$$

**F1 Score**:

The F1 Score could be a more appropriate metric to consider when there is a need to strike a balance between Precision and Recall, F1 score is the harmonic mean of Precision and Recall and is high if both of the values are well balanced

$$\text{F1} = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

### 2.2 Efforts on Evaluation Standardization

With time some organizations have been putting effort to standardize the method by which the evaluation of Sentiment analysis models is verified. This is being done as in the current scenario the models being made are not being set on an equal footing to not only be able to objectively judge their performance against each other but also understand where the strengths of a certain model lie. While it would be a boon to have a system that performs equally in all cases it is not realistic. Thus, we need a testing mechanism that not only gives an idea about the model or algorithm or approaches objective performance but is also able to reveal what the model is best for.

One such organization is The International Workshop of Semantic Evaluation (IWSE). With their work, they focus on different aspects of Semantic analysis and evaluation of different ideas and algorithms. They provide each competing party the same unannotated data and allow them to detect the Aspect and Sentiment in the same and then each work is evaluated against a fair system that results in an overview of each system that can be compared on equal footing.

The issues faced with such evaluation standardization procedure are as follows:

- Diversity of application and requirements – With the vast applications that entail Sentiment analysis research there are bound to be points where one type of evaluation will not be a good representation of the other due to their conflicting process.
- Different annotation techniques – the procedure of annotation itself is affected by inconsistency and standardizing one over the other will cause us to misfit a use case. For example, the annotation of sentiment polarity is limiting for a use case that needs a system that uses a 5-star annotation method.
- The need to understand what a task needs – coming back to vast applications, it is essential that we expand in both the depth of evaluation and the breadth of applications which needs a shifting focus.

## 3 Core Solutions

This chapter dives into the papers discussing their intuition and ideas. The idea is as follows: First, we discuss the intuition of the authors in the paper, Next, we discuss the model architecture they use to perform the task and finally, we talk about the evaluation datasets used by the authors.

The entire performance is discussed in three tables that talk about the multilingual performance for papers 3.2 and 3.3, The BERT and non-BERT performance based on 3.5 and 3.6, and the overall performances of the models.

**3.1 Triplet Extraction [7]**

An aspect Sentiment triplet can be defined as *(aspect, opinion, sentiment)* Where the aspect is the target word about which the triplet is built, the sentiment is the sentimental polarity that is given to the aspect and the opinion is the words that were the emotional expression of the aspect term. These triplets allow us to capture more emotional factors of the sentence and make it more usable in practical scenarios.

**The model**

The novel idea presented by the author is:

A multi-task dual encoder using the BERT token and Label Encoders to build tokens for the text and part of speech in the text, which is then passed through two sets of grids that can map the relationships between the aspects and opinion terms and make boundaries around these terms. The intuition appeared from the idea that the former techniques ignored the major part of the text i.e., the relationship between these words, like the one-to-many, many-to-one, etc. that better clarified the ideas in these sentences. The first grid does the major lifting and the second grid performs the regularization of these triplets.

The model also uses a new ten-tags system to better label the label tokens. These tags allow the inference in the following ways – It makes it easier for the heuristic of the model to infer if the aspect word is single and strengthens the heuristic by tags in case the aspect word is multiple due to its rules. Also, tagging one pair of words allows the combination of those words with other words to be more predictable as even with a unique relationship it tends to follow a pattern.

**Evaluation**

The datasets from [8] and [9] were derived from SemEval14 [5], SemEval16[7], and SemEval15. The model outperforms the baseline models in all cases. When compared to the GTS-BERT framework the model showed an increase in F1-Score by a percent of 2.03%, 3.9%, 5.57%, and 2.41% on 14res, 14lap, 15res, and 16res datasets respectively.

**3.2 Local Context and Global Context in Aspect-Based Sentiment Multilingual Analysis [15]**

In most methods for aspect-level sentiment analysis, the local context is considered the most due to the thought that the local context would hold the most relevant contextual info. However, the global context has a deep connection to aspects of local sentiment and should be taken into consideration. To focus on both the local and global context of the sentence as the sentiment polarity of an aspect is based on both. The Local Context Focus is derived using multi-headed self-attention (MHSA) and Context-features Dynamic Mask (CDM) or Context-features Dynamic Weighting (CDW) [15] and Global Context Focus is derived using a BiGRU, CNN, and Layer Normalization. All of this is done to make a step towards building a system that can work on a multilingual learning model by using both Chinese and English data sets.

**Model**

Using the BERT model [11], which is pre-trained to generate the contextual word embeddings that can, we generate word embeddings for both branches

$$O^{l}_{BERT} = BERT^{l}\left(X^{l}\right)$$
$$O^{g}_{BERT} = BERT^{g}\left(X^{g}\right)$$

$O^{l}_{BERT}$ and $O^{g}_{BERT}$ are the BERT-shared embedding layers for both contexts. As discussed formerly, the local context focus uses a multi-headed self-attention to generate attention matrices that will dictate the focus on the context words for any aspect or sentence. The concerning local words are detected using the method to find the relative semantic distance. The Context features that are not learnable with high recall by Bert sharing layer are masked by the CDM layer to minimize the variance they will bring without losing information by removing them entirely and the CDW layer would dynamically weigh down semantically distant features from the aspect to allow for a higher focus on the local context.

To go with these, A BiGRU is used with tanh as an activation function. The input is passed

[Type here]
Journal of Current Research in Engineering and Science
Bi-Annual Online Journal ( ISSN : 2581 - 611X )

Volume 6- Issue 2, August 2023
Paper : 34

through a CNN to pop out the important features in the global context. Finally, a layer normalization (LN) is applied that can help the model avoid problems that come with mini-batch distribution in batch normalization, making it viable for dynamic network scenarios, RNN, etc. The next layer uses the output to learn the global context feature. Finally, the output yields feature representations divided by the rest tokens

**Evaluation**

The model evaluation was done on three Chinese datasets [12]. The English datasets were SemEval-2014 Laptop, Restaurant datasets [5], the SemEval-2016 Restaurant dataset [6], an ACL Twitter social dataset [13], the T-shirt dataset, and the Television dataset [14]. Some variations of the model (LGCF-CDM, LGCF-CDW, LGCF-CDM-CDW) reigned better than all other baseline models for all three datasets, providing a higher Accuracy and F1-Score by far with only the BERT-Base model coming any close.

### 3.3 Using a Separate Knowledge Base to Assist a GCN in Chinese-oriented Aspect-Based Sentiment Analysis [16]

The Proposed model uses the word embeddings generated from BERT and uses SenticNet as the Semantic knowledge base [16] to solve the problem of most neural networks depending on syntactic dependencies and ignore the semantic commonsense knowledge. Also, the model performs the task of sentiment analysis for multilingual datasets. The model uses the relationship between nodes of the GCN to represent the relationship between sentences providing relative context and aspect information. SenticNet is a public semantic resource *( https://sentic.net/)*. And it is used as a dictionary for the proposed model.

**Model**

The model is built upon a GCN that uses words as nodes with the aspect words made into more weighted nodes. Some aspects are connected via inter-aspect connection as they need the other in sentiment classification. Finally, the dependency graph amongst the words is changed using the knowledge base from SenticNet. The adjacency matrix from the dependency graph and SenticNet is used to train the GCN layer.

**Evaluation**

The model has beaten the baseline models in all of the datasets from the two languages except for the Car dataset from the Chinese dataset and The Rest14 [5] from the English dataset. The model outperformed in the case of Macro-F1 score when compared to the second highest in the cases of Rest15, 16, and MAMS dataset by 1.15, 5.70, and 1.39 percent. This was possible due to the external knowledge provided to the system via the SenticNet Knowledge Base.

### 3.4 A multi-level architecture using BERT, BiLSTM, GCN, and CNN to find hidden aspects [19]

The following model proposes an approach to use a CNN model laid over a BERT-GCN model to tackle two major problems faced by GCN – Limited layers due to Vanishing Gradient Problem and the inability to asses hidden contexts. For ex, “The restaurant has six different types of desserts” should be classified as positive. Rather it is classified as neutral due to a lack of context words.

**Model**

The paper proposes a novel model that uses CNN in combination with BERT GCN and BiLSTM. The function of each layer is explained as follows:
BERT - Generates embeddings using attention models. These word embeddings are contextually sound and are helpful in the inference of aspect and Sentiment connections.
BiLSTM - Creates Contextualized Word Representations. These word representations are generated from the word embeddings from the Bert layer. Due to its Bidirectional nature, the BiLSTM model can generate a representation of the word that is relevant to the sentence as a whole.
GCN - Extracts significant Features over the Contextualized word Representations as these word representations contain useless words that are still not very useful to the model thus the GCN using its dependency graph can extract exactly what part of the sentence is essential for us.

CNN - Performs Sentiment Analysis on these feature vectors from GCN. The CNN layer solves the problem faced by the standard GCN layer of not being able to go too deep due to vanishing and exploding gradient problems.

**Evaluation**

Using Three Benchmark datasets the Laptop, Restaurant [5], and Twitter(http://goo.gl/5Enpu7) to generate the following results For the three datasets the F1 score is the lowest for the Twitter Dataset and highest for the Restaurant dataset even though the Twitter dataset is twice the size of the Restaurant data. This can be attributed to the issue of data quality. Thus, this shows a shortcoming of the dataset that it requires better-pre-processed data. Otherwise, the performance of the model will suffer. The strong point of the model is that it uses the best qualities of each of its components to provide better predictions.

### 3.5 Using a syntax tree to model Aspect based Sentiment analysis model [20]

With the usage of neural networks for Sentiment analysis we face the problem of effectively modeling long-term dependencies in Aspect-Level Sentiment detection and have a problem when the sentences have multiple aspects in them. The author suggests the idea to use a syntax graph made from a syntax tree (a tree that has a root node as a phrase or a word that is then connected to the rest of the words of the sentence based on its relationship to them) that has the aspect term defined as the root point and connects to the rest of terms based on relationships.

**Model**

Specifically, the proposed model known as the RSSG aka Reliable Search on Syntax Graph assigns weight to context words depending on their qualitative and quantitative relationship with the aspect. We further decrease the parsing error due to incorrect syntactic dependencies based on using a convolution layer to adjust the aspect-dependent weights and extract strongly aspect-related context words. The proposed model focuses on each aspect word separately and follows the following procedure:

- We build a syntax tree that is rooted in the aspect word where each of the other words is connected to it directly or indirectly to represent a syntactic dependency. For a multiword aspect, the aspect's headword is considered the tree root.
- We build a syntax graph from the tree in which the following rules are followed – a self-loop at each node, each word node is connected to the aspect node via a directed edge from the aspect node to the word node.
- Generate and concatenate the word embeddings, POS embeddings, and syntax-based position embeddings, to extract the input embeddings for each word using the GLoVe [26] word embedding. Besides this, the same work is done with BERT-based Embeddings and compared against BERT-based models.
- Use a GRU layer to build contextual representation out of the word embeddings.
- Syntax-guided searching – A breadth-first searching method starting from the aspect node assigns weights to the terms based on their word embeddings, contextual representations, and distance from the aspect node to understand their relevance.
- Employ a convolutional layer to assemble features that capture different lengths of term dependencies.
- Classifier – We use a Dense layer with SoftMax activation to get predictions.

**Evaluation**

The paper evaluated the model on four benchmark datasets – Res14, Lap14 from SemEval2014 [5], Twitter [13] and the following are their performance compared to the baseline models. The model uses two-word embeddings – GLoVe [26] and BERT generated. As we can see, the BERT-based RSSG showed a significant increase in its scores owning to the better-contextualized word embeddings that BERT can generate compared to GLoVe.

### 3.6 An end-to-end Aspect Based Sentiment analysis model that uses the Syntactic

**structure and the Semantic information from the Lexical [21]**

The paper discusses two major problems seen while using Graph Neural Networks - in the existing models that use end-to-end architecture, the use of dependency tree and GNN is rarely seen and irrespective of using GCN [17] or GAT [18] to the dependency tree, the information that is derived from established connections between tree node and other nodes is used neglecting specific nodal connections leading to incomplete use of the tree, and the syntactic structures by GNN from a syntactic tree are also incomplete. Also, instead of modelling the words, the author bases their modelling on the smallest semantic unit i.e., sememes give a better understanding of the semantic information. Sememe as defined by Google is the unit meaning carried by morphemes which are defined as a meaningful morphological word that cannot be divided further.

**Model**

To solve these issues Y. Bie et al. [21] propose a novel model based on fusing the Syntactic information and lexical information (SSi-LSi) that is made of two branches after the embedding layer. The first branch uses an improved relation-attention GCN, which processes the dependency tree to extract syntactic information. The other Branch uses a BiGRU to encode embeddings that are contextual and uses another method improved from [25] to generate word representations based on Part of Speech information and lexical sememes and integrate the output from both using an attention mechanism. Finally, the output of the two branches is fused and passed through the decoding layer to generate outputs.

**Evaluation**

Datasets for evaluations – Laptop dataset, Restaurant dataset from SemEval-2014 [5], and English Tweets [23]. While comparable to most of the pipeline base models in general, the model has an important advantage over these models as it doesn't suffer from the standard problems of the pipeline models like error propagation and accumulation. It also outperforms the models like INABSA and MNN model which also follows the similar two-way architecture of the author's model. Also, the model performs way better when combinedly used with BERT in place of the embedding layer. BERT allows the model to fine-tune better as is more versed in the specific task.

**3.7 Aspect-Level Sentiment Classification Based on Auto-Adaptive Model Transfer [22]**

Uses an auto-adaptive model transfer to transfer learned parameters from a document-based Sentiment analysis model to Aspect based sentiment analysis model.
Targets the lack of Aspect level Sentiment analysis datasets and the comparative abundance of the DLSC datasets and uses transfer learning from document-level Sentiment analysis to Aspect level Sentiment analysis. The Auto Adaptive Model transfer will adapt to the changes in domain and task differences. Using disturbance variables that are selected based on the task allows us to capture the unique part that distinguishes tasks and also use an attention mechanism to build vocabulary

**Model**
The Baseline model contains an embedding for word embeddings. The word embedding is sent to a BiLSTM layer that obtains context information about the sentence. Next, we introduce an attention layer for target aspects. This layer uses the aspects to compute the attention weights for the target and generate better sentiment vocabulary. Lastly, the model uses a fully connected output layer using the SoftMax function to generate outputs.
The Auto-Adaptive model Transfer method uses the idea of separate modules and separate parameters with some shared parameters that are concatenated with the disturbance variables for each of the tasks. Thus, it differs from Vanilla Model Transfer as it assumes that the shared parameters between the two tasks are the same. The loss function is joined over all the parameters and allows us to learn all the parameters for each separate module, shared module, and the disturbance variables at the same time.
**Evaluation**
Dataset used – Res14, Laptop14[5], and Res16 [24]. The model shows better metrics than all baseline models owning to its combined

[Type here]
Journal of Current Research in Engineering and Science
Bi-Annual Online Journal ( ISSN : 2581 - 611X )

Volume 6- Issue 2, August 2023
Paper : 34

learning power to learn global and local contextual words using document-based and aspect-based analysis. The model only falters in accuracy in the Rest16 dataset but has shown an overall better F1 score than all the other models suggesting a good balance of recall and Precision.

**3.8 Overall Evaluation of the Models**

All the models were tested commonly on the SemEval14 Restaurant dataset [15]. Thus, if not specified the performance is for that dataset. Table 1-3

| Paper | Model | English dataset performance | Non-English dataset performance |
|---|---|---|---|
| **J. He et al. [15]** | Two-way model sharing a BERT-based embedding layer | Acc = 85.52<br>F1-score = 79.85 | Camera –<br>Acc = 97.24<br>F1-score = 96.39<br>Phone –<br>Acc = 97.59<br>F1-score = 97.17<br>Car –<br>Acc = 98.36<br>F1-score = 97.89 |
| **Q. Yang et al. [16]** | GCN using SenticNet as a knowledge base | Acc = 86.79<br>F1-score = 81.03 | Camera –<br>Acc = 97.71<br>F1-score = 97.05<br>Phone –<br>Acc = 97.43<br>F1-score = 97.04<br>Car –<br>Acc = 97.40<br>F1-score = 96.82 |

Table 1: Multilingual analysis

| Paper | Model | BERT | Non-BERT |
|---|---|---|---|
| **Y. Bie et al. [21]** | Two-Branch model using BiGRU and GCN | Acc = NA<br>F1 = 70.13 | Acc = NA<br>F1 = 67.28 |
| **R. Zhang et al. [20]** | A GRU and CNN-based model using syntax tree | Acc = 87.0<br>F1 = 81.3 | Acc = 84.2<br>F1 = 77.0 |

Table 2: Comparison of models from papers 3.5 and 3.6 using BERT as the embedding layer the BERT vs the non-BERT embedding layer

[Type here]
Journal of Current Research in Engineering and Science
Bi-Annual Online Journal ( ISSN : 2581 - 611X )

Volume 6- Issue 2, August 2023
Paper : 34

| Paper | Focus | Model | Performance |
|---|---|---|---|
| **H. Huan et al. [7]** | Aspect-based Triplet extraction | Dual-Encoder | Acc = NA<br>precision = 74.12<br>Recall=72.84<br>F1 = 73.47 |
| **J. He et al. [15]** | Multilingual Sentiment analysis using Local context and global context | Two-way model sharing a BERT-based embedding layer | Acc = 85.52<br>F1 = 79.85 |
| **Q. Yang et al. [16]** | Using a knowledge base (SenticNet) to assist the process of aspect-detection using GCN | GCN using SenticNet as a knowledge base | Acc = 86.79<br>F1 = 81.03 |
| **H. T. Phan et al. [19]** | Using Multilevel specialized architecture to build a deeper and more robust model | BERT-BiLSTM with GCN and CNN | Acc= 85.25<br>F1 = 78.76 |
| **R. Zhang et al. [20]** | Use a Syntax tree-based model that can use syntactic and semantic hints to better perform sentiment analysis | A GRU and CNN-based model using syntax tree | Acc = 87.0<br>F1 = 81.3 |
| **Y. Bie et al. [21]** | Use both the semantic and Lexical information (sememes) as the base modeling unit. | Two-Branch model using BiGRU and GCN | Acc = NA<br>F1 = 70.13 |
| **W. Zheng et al. [22]** | Use a DLSC dataset to train a baseline model and use its training parameters to help in Transfer Learning an ALSC model | Auto-Adaptive model transfer | Acc = 82.5 F1 = 74.28 |

Table 3: Model Performance comparison. Only the best performance is considered

# 4 Challenges and Conclusions

### 4.1 Challenges

The linguistic task of automatically detecting aspects and related sentiments is quite challenging due to complex linguistic phenomena that are difficult to interpret and understand. Identifying the appropriate aspect can be considered the most difficult part of this analysis. Nevertheless, correctly identifying the sentiment related to a particular aspect can be challenging. Here are some of the main challenges in ABSA:

- **Aspect identification**: The first challenge in ABSA is to identify the aspects or entities that are being talked about in the text. This can be difficult because aspects can be ambiguous and can be expressed in different forms (e.g., single words, noun phrases, verb phrases).

- **Aspect categorization**: Once the aspects have been identified, the next challenge is to categorize them into different sentiment categories (e.g., positive, negative, neutral). This can be challenging because an aspect can have multiple sentiment orientations based on its context.
- **Contextual understanding**: The sentiment orientation of an aspect can be influenced by the context in which it appears. Therefore, it is important to understand the context of the aspect and its relation to other aspects of the text.
- **Data sparsity**: ABSA requires large amounts of labeled data for training machine learning models. However, it can be challenging to obtain labeled data for all possible aspects and sentiment categories, leading to data sparsity.
- **Domain adaptation**: Sentiment expressions can vary across different domains and contexts. Therefore, it can be challenging to develop a model that performs well across different domains.
- **Negation and sarcasm**: Negation and
- sarcasm can flip the polarity of a sentiment expression, making it difficult to accurately identify the sentiment orientation of an aspect.
- **Multi-lingual and cross-lingual sentiment analysis**: ABSA is more challenging for languages other than English, as it requires domain-specific knowledge, context, and sentiment resources that are not always available in other languages.

These challenges make ABSA a complex and dynamic research area, requiring innovative techniques and approaches to overcome them.

### 4.2 Conclusion

With this final note, we end the survey. The survey focused on multiple methodologies of Sentiment analysis that changed both the core system of the classifier and the method in which the data is processed or how different methods are used to support the model to make better solutions. We discovered how different technologies can be stacked or used one over the other to build better solutions that overcome the problems of a single technology. We also discussed the current efforts being done to standardize different aspects of Sentiment analysis and Natural Language Processing as a whole and what issues such efforts are facing.

## REFERENCES:

[1]. B. Liu, Sentiment Analysis and Opinion Mining (series Synthesis Lectures on Human Language Technologies). vol. 16. San Mateo, CA, USA: Morgan, 2012.

[2]. M. Tsytsarau and T. Palpanas, “Survey on mining subjective data on the web,” *Data Mining Knowl. Discovery*, vol. 24, no. 3, pp. 478–514, 2012.

[3]. K. Schouten and F. Frasincar, "Survey on Aspect-Level Sentiment Analysis," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813-830, 1 March 2016, doi: 10.1109/TKDE.2015.2485209.

[4]. Nazir, Y. Rao, L. Wu and L. Sun, "Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey," in *IEEE Transactions on Affective Computing*, vol. 13, no. 2, pp. 845-863, 1 April-June 2022, doi: 10.1109/TAFFC.2020.2970399.

[5]. M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, SemEval-2014 task 4: Aspect-based sentiment analysis, in

Proc. 8th Int. Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 2014, pp. 27–35.

[6]. M. Pontiki et al., "Semeval-2016 task 5: Aspect-based sentiment analysis," in Proc. SemEval, 2016, pp. 19–30.

[7]. H. Huan, Z. He, Y. Xie, and Z. Guo, "A Multi-Task Dual-Encoder Framework for Aspect Sentiment Triplet Extraction," in *IEEE Access*, vol. 10, pp. 103187-103199, 2022, doi: 10.1109/ACCESS.2022.3210180.

[8]. H. Peng, L. Xu, L. Bing, F. Huang,W. Lu, and L. Si, ``Knowing what, how and why: A near complete solution for aspect-based sentiment analysis," in Proc. AAAI Conf. Artif. Intell., vol. 34, no. 5, Apr. 2020, pp. 8600_8607.

[9]. L. Xu, H. Li, W. Lu, and L. Bing, ``Position-aware tagging for aspect sentiment triplet extraction," in Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP), 2020, pp. 2339_2349.

[10]. Z. Wu, C. Ying, F. Zhao, Z. Fan, X. Dai, and R. Xia, ``Grid tagging scheme for aspect-oriented _ne-grained opinion extraction," in Proc. Findings Assoc. Comput. Linguistics, (EMNLP), 2020, pp. 2576_2585.

[11]. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, ``BERT: Pretraining of deep bidirectional transformers for language understanding," in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol., vol. 1. Minneapolis, MN, USA: Association for Computational Linguistics, Jun. 2019, pp. 4171_4186. [Online]. Available: https://www.aclweb.org/anthology/N19-1423, doi: 10.18653/v1/N19-1423

[12]. H. Peng, Y. Ma, Y. Li, and E. Cambria, ``Learning multi-grained aspect target sequence for Chinese sentiment analysis," Knowl.-Based Syst., vol. 148, pp. 167_176, May 2018, doi: 10.1016/j.knosys.2018.02.034.

[13]. L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, ``Adaptive recursive neural network for target-dependent Twitter sentiment classi_cation," in Proc. 52nd Annu. Meeting Assoc. Comput. Linguistics, vol. 2, 2014, pp. 49_54, doi: 10.3115/v1/p14-2009.

[14]. R. Mukherjee, S. Shetty, S. Chattopadhyay, S. Maji, S. Datta, and P. Goyal, ``Reproducibility, replicability and beyond Assessing production readiness of aspect-based sentiment analysis in the wild," in Proc. Eur. Conf. Inf. Retr., in Lecture Notes in Computer Science, vol. 12657, D. Hiemstra, M.-F. Moens, J. Mothe, R. Perego, M. Potthast, and F. Sebastiani, Eds. Springer, Mar./Apr. 2021, pp. 92106. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-72240-1_7, doi 10.1007/978-3-030-72240-1_7.

[15]. J. He, A. Wumaier, Z. Kadeer, W. Sun, X. Xin and L. Zheng, "A Local and Global Context Focus Multilingual Learning Model for Aspect-Based Sentiment Analysis," in *IEEE Access*, vol. 10, pp. 84135-84146, 2022, doi: 10.1109/ACCESS.2022.3197218.

[16]. Q. Yang, Z. Kadeer, W. Gu, W. Sun and A. Wumaier, "Affective Knowledge Augmented Interactive Graph Convolutional Network for Chinese-Oriented Aspect-Based Sentiment Analysis," in *IEEE Access*, vol. 10, pp. 130686-130698, 2022, doi: 10.1109/ACCESS.2022.3228299.

[17]. T. N. Kipf and M. Welling, Semi-supervised classification with graph convolutional networks, arXiv preprint arXiv: 1609.02907, 2016.

[18]. P. Veliˇckovi´c, G. Cucurull, A. Casanova, A. Romero, P. Li`o, and Y. Bengio, Graph attention networks, arXiv preprint arXiv: 1710.10903, 2017.

[19]. H. T. Phan, N. T. Nguyen, and D. Hwang, "Aspect-Level Sentiment Analysis Using CNN Over BERT-GCN," in *IEEE Access*, vol. 10, pp. 110402-110409, 2022, doi: 10.1109/ACCESS.2022.3214233.

[20]. R. Zhang, Q. Chen, Y. Zheng, S. Mensah, and Y. Mao, "Aspect-Level Sentiment Analysis via a Syntax-Based Neural Network," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 30, pp. 2568-2583, 2022, doi: 10.1109/TASLP.2022.3190731.

Volume 6- Issue 2, August 2023
Paper : 34

[21]. Y. Bie, Y. Yang, and Y. Zhang, "Fusing Syntactic Structure Information and Lexical Semantic Information for End-to-End Aspect-Based Sentiment Analysis," in *Tsinghua Science and Technology*, vol. 28, no. 2, pp. 230-243, April 2023, doi: 10.26599/TST.2021.9010095.

[22]. W. Zheng, H. Jin, Y. Zhang, X. Fu and X. Tao, "Aspect-Level Sentiment Classification Based on Auto-Adaptive Model Transfer," in *IEEE Access*, vol. 11, pp. 34990-34998, 2023, doi: 10.1109/ACCESS.2023.3265473.

[23]. M. Mitchell, J. Aguilar, T. Wilson, and B. Van Durme, Open domain targeted sentiment, in Proc. 2013 Conf. Empirical Methods in Natural Language Proc., Seattle, WA, USA, 2013, pp. 1643–1654.

[24]. M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, ‘‘SemEval-2016 task 5: Aspect based sentiment analysis,’’ in Proc. 9th Int. Workshop Semantic Eval. (SemEval), 2016, pp. 19–30.

[25]. L. Luo, X. Ao, Y. Song, J. Li, X. Yang, Q. He, and D. Yu, Unsupervised Neural Aspect Extraction with Sememes, in Proc. 28th Int. Joint Conf. Artificial Intelligence, Macao, China, 2019, pp. 5123–5129.

[26]. Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation